

25.06.99

日 本 国 特 許 庁
PATENT OFFICE
JAPANESE GOVERNMENT

REC'D 13 AUG 1999

WIPO PCT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application:

1999年 2月17日

出 願 番 号
Application Number:

平成11年特許願第039253号

出 願 人
Applicant(s):

松下電器産業株式会社

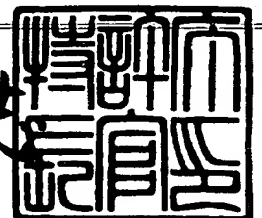
PRIORITY
DOCUMENT

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

1999年 7月15日

特許庁長官
Commissioner,
Patent Office

山 建 志 郎



出証番号 出証特平11-3050213

【書類名】 特許願

【整理番号】 2015200158

【提出日】 平成11年 2月17日

【あて先】 特許庁長官殿

【国際特許分類】 G10L 3/00
G06F 15/38

【発明者】

【住所又は居所】 大阪府門真市大字門真 1 0 0 6 番地 松下電器産業株式
会社内

【氏名】 脇田 由実

【特許出願人】

【識別番号】 000005821

【氏名又は名称】 松下電器産業株式会社

【代理人】

【識別番号】 100092794

【弁理士】

【氏名又は名称】 松田 正道

【電話番号】 066397-2840

【手数料の表示】

【予納台帳番号】 009896

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9006027

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 言語変換装置

【特許請求の範囲】

【請求項 1】 ある文とその文を異なる言語または文体などに翻訳した文（以下、目的言語文と呼ぶ、これに対応して翻訳する対象となったもとの「ある文」を原言語文と呼ぶ）とを一对とした文データベース（以下、対訳コーパスという）と、

その文データベース中の原言語文及び目的言語文における単語または品詞の隣接頻度を算出し、頻度の高い単語及び品詞を連結して意味的なまとまりを形成する部分文（以下、フレーズと呼ぶ）を抽出するフレーズ抽出部と、

前記フレーズ抽出部で抽出された前記フレーズで、原言語及び目的言語のフレーズの間接関係を調べることで対応するフレーズを決定するフレーズ決定部と、

決定された前記対応するフレーズを保管しておくフレーズ辞書と、

原言語文が入力された際にこの入力文と前記フレーズ辞書に格納されている前記対応するフレーズとを照合することで言語または文体変換を行う言語変換部とを備えたことを特徴とする言語変換装置。

【請求項 2】 前記フレーズ決定部は、原言語及び目的言語のフレーズの間接関係を調べることで対応するフレーズを決定することを特徴とする請求項 1 記載の言語変換装置。

【請求項 3】 前記対訳コーパスの原言語文を単語列に変換する形態素解析部と、

その形態素解析部の結果を利用して原言語文及び目的言語文の一部または全部の単語を品詞名で置き換えた対訳コーパスを作成する品詞化部を更に有し、

前記フレーズ抽出部は、前記品詞化部で品詞化された対訳コーパスからフレーズを抽出することを特徴とする請求項 1 記載の言語変換装置。

【請求項 4】 原言語と目的言語との対訳単語辞書を有し、

前記品詞化部は、前記対訳単語辞書で対応付けされている単語でかつ原言語が内容語である単語を品詞化することを特徴とする請求項 3 記載の言語変換装置。

【請求項 5】 前記フレーズ抽出部は、予め優先的にフレーズとみなしたい単

語または品詞列を原言語と目的言語を対にして保管しておくフレーズ定義表をも利用して、フレーズを抽出することを特徴とする請求項 1 記載の言語変換装置。

【請求項 6】 請求項 1 ～ 5 のいずれかに記載の言語変換装置の各構成要素の全部または一部の機能をコンピュータに実行するためのプログラムを格納していることを特徴とするプログラム記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、入力音声または入力テキストを、他言語または他の文体型などに変換して出力する言語変換装置に関するものである。

【0002】

【従来の技術】

以下、従来の技術を言語変換装置の 1 つである、入力音声了他言語に翻訳（以下通訳と呼ぶ）する通訳装置を例にして説明する。以下通訳装置における原体型文を原言語文、目的体型文を目的言語文と呼ぶ。

【0003】

通訳装置は、音響信号として入力された発声文を単語テキスト列で表示された出力文に変換するための音声認識と、単語テキスト列で表示された文を入力し他言語文に翻訳する言語翻訳とを順次実行することで通訳を実現している。さらに上記言語翻訳部は、入力文の統語的または意味的構造を解析する言語解析部と、解析結果に基づいて他言語に変換する言語変換部と、翻訳結果から自然な出力文を生成する出力文生成部とから構成されている。

【0004】

しかし、音声認識部が発声文の一部を誤認識した場合や、文にあいづちや言い直しなどが挿入されたり、一部の単語が省略されて話された場合など、文自体が統語的に不自然な場合は、音声認識結果を言語解析部に入力しても解析が失敗し、結果的に翻訳結果が出力されないという問題があった。

【0005】

この問題を解決するために、従来の文法に則って言語解析を行うのではなく、

従来の文法では解析できないような発声文も含めた発声文例から発声文特有の文法を作成し、作成された文法を用いて言語解析と言語変換とを行う方法が提案されている。(たとえば、古瀬、隅田、飯田：情報処理学会論文誌Vol35,no3,1994-3)。以下、文献の従来例を図6の通訳装置例に沿って説明する。

【0006】

まず通訳を行う前に、予め発声文対訳コーパスから発声文規則を作成する。ここでは、一部の単語が誤ったり省略されたりすることを考慮し、発声文例を意味的なまとまりとしての最小単位(以後、フレーズと呼ぶ)に分割し、フレーズ内規則とフレーズ間の依存規則とを作成している。まずフレーズ決定部61で、対訳発声文例をフレーズ単位に分割し、フレーズ内規則とフレーズ間の依存関係規則を作成する。たとえば、「部屋の予約をお願いしたいんですが」「I'd like to reserve a room」という対訳文例は、(a)「部屋の予約」「reserve a room」、(b)「をお願いしたいんですが」「I'd like to」という(a)(b)2つの対訳フレーズに分割され、「(a)を(b)する」「(b) to (a)」という依存関係が規則化される。上記対訳フレーズは対訳フレーズ辞書62に、フレーズ間の依存関係を対訳の形で表されたものはフレーズ間規則テーブル63に各々保管される。このような処理が対訳コーパスに含まれた全発声文分について行われる。このフレーズの分割と依存関係は、文の意味的情報やどの程度文法的に崩れていないかの度合いなどのファクターから決定されるため、自動的に各文について決定することが難しく、従来は人手で決定されている。

【0007】

次に通訳の際には、まず発声された原言語音声は音声認識部64に入力される。音声認識部では、たとえば、対訳フレーズ辞書62にフレーズとして記述されている単語のネットワークとフレーズ間規則テーブル63にて記述されている隣接フレーズbi-gramとにより、時系列に沿って順次認識単語候補が予測される。

予め学習されている音響モデル13と入力音声との距離値をベースとした音響スコアとフレーズbi-gramによる言語スコアとの和を認識スコアとし、Nbest-searchにより認識候補である連続単語列が決定される。

【0008】

認識された連続単語列は、言語変換部 65 に入力される。言語変換部では、入力された連続単語列を対訳フレーズ辞書 62 内のフレーズ列に変換し、各フレーズ列に相当するフレーズ間規則を探索する。そして、各フレーズの対訳である目的言語フレーズと目的言語のフレーズ間規則とから、入力原言語認識結果文を目的言語文に変換する。

【0009】

変換された目的言語文は出力文生成部 66 に入力され、統語的な不自然さを修正する。たとえば、定冠詞や不定冠詞の付与、代名詞、動詞における 3 人称化や複数化や過去形化などの最適化などが行われる。修正後の目的言語翻訳結果文はたとえばテキストとして出力される。

【0010】

【発明が解決しようとする課題】

しかしながら、従来例における言語変換装置である通訳装置に用いた対訳フレーズ辞書やフレーズ間規則テーブルは、発声文の意味的情報や文法的情報を自動的に解析できる手段がないために、人手で作成しなければならない。そのため、開発に時間がかかり、人手の作成基準の揺れが規則性能を歪ませるという問題点がある。たとえば、通訳装置の目標となるタスクを変更したり、原言語及び目的言語の言語種が変更になった場合は、一度構築した規則を適応できずにはじめから規則を作成しなければならず、開発効率が悪く手間がかかる、という課題を有していた。

【0011】

本発明は、対訳フレーズ辞書やフレーズ間規則テーブルは人手で作成しなければならず、開発効率が悪く人手の作成基準の揺れが規則性能を歪ませるという課題を考慮し、対訳フレーズ辞書作成や対訳フレーズ間規則テーブルの作成をなるべく自動的に行うことを可能とすることで、効率よく品質の高い対訳フレーズ辞書を生成できる言語変換装置を提供することを目的とするものである。

【0012】

【課題を解決するための手段】

上述した課題を解決するために第 1 の本発明（請求項 1 に対応）は、ある文と

その文を異なる言語または文体などに翻訳した文（以下、目的言語文と呼ぶ、これに対応して翻訳する対象となったもとの「ある文」を原言語文と呼ぶ）とを一对とした文データベース（以下、対訳コーパスという）と、

その文データベース中の原言語文及び目的言語文における単語または品詞の隣接頻度を算出し、頻度の高い単語及び品詞を連結して意味的なまとまりを形成する部分文（以下、フレーズと呼ぶ）を抽出するフレーズ抽出部と、

前記フレーズ抽出部で抽出された前記フレーズで、原言語及び目的言語のフレーズの間接関係を調べることで対応するフレーズを決定するフレーズ決定部と、決定された前記対応するフレーズを保管しておくフレーズ辞書と、

原言語文が入力された際にこの入力文と前記フレーズ辞書に格納されている前記対応するフレーズとを照合することで言語または文体変換を行う言語変換部とを備えたことを特徴とする言語変換装置である。

【0013】

また第2の本発明（請求項2に対応）は、前記フレーズ決定部は、原言語及び目的言語のフレーズの間接関係を調べることで対応するフレーズを決定することを特徴とする第1の発明に記載の言語変換装置である。

【0014】

また第3の本発明（請求項3に対応）は、前記対訳コーパスの原言語文を単語列に変換する形態素解析部と、

その形態素解析部の結果を利用して原言語文及び目的言語文の一部または全部の単語を品詞名で置き換えた対訳コーパスを作成する品詞化部を更に有し、

前記対訳フレーズ抽出部は、前記品詞化部で品詞化された対訳コーパスからフレーズを抽出することを特徴とする第1の発明に記載の言語変換装置である。

【0015】

また第4の本発明（請求項4に対応）は、原言語と目的言語との対訳単語辞書を有し、

前記品詞化部は、前記対訳単語辞書で対応付けされている単語でかつ原言語が内容語である単語を品詞化することを特徴とする第3の発明に記載の言語変換装置である。

【0016】

また第5の本発明（請求項5に対応）は、前記フレーズ抽出部は、予め優先的にフレーズとみなしたい単語または品詞列を原言語と目的言語を対にして保管しておくフレーズ定義表をも利用して、フレーズを抽出することを特徴とする第1の発明に記載の言語変換装置である。

【0017】

次に本発明の動作を説明する。

【0018】

請求項1に記載の言語変換装置においては、隣接する頻度の高い単語または品詞列を連結してフレーズを生成し、原言語文及び目的言語文を対応させながら対訳フレーズを決定することにより、コーパスから対訳フレーズを自動的に抽出することを可能とし、人手をなるべく用いずに、効率よく品質の高い対訳フレーズ辞書を生成できる。

【0019】

また、請求項2に記載の言語変換装置は、対訳フレーズを決定する際に、原言語及び目的言語のフレーズの共起頻度を調べることで、コーパスから対訳フレーズを自動的に抽出することを可能とし、人手をなるべく用いずに、効率よく品質の高い対訳フレーズ辞書を生成できる。

【0020】

また、請求項3及び4に記載の言語変換装置においては、予め対訳コーパスの一部の単語を品詞化しておき、品詞化された対訳コーパスから対訳フレーズを抽出することで、自動的にフレーズを抽出する際に、コーパスに含まれる単語以外の単語が含まれているフレーズにも対応できるため、より効率よく品質の高い対訳フレーズ辞書を生成できる。

【0021】

また、請求項5に記載の言語変換装置は、請求項1の動作で自動抽出される対訳フレーズと、予め人為的に定義された対訳フレーズとの両方を対訳フレーズ辞書に登録することにより、人手を最小限度に押さえた範囲で、より品質の高い対訳フレーズ辞書を生成できる。

【0022】

【発明の実施の形態】

以下に本発明の実施の形態について図面を参照して説明する。

【0023】

(実施の形態1)

以下、図面を参照して本発明の請求項1から4の実施の形態を説明する。本実施の形態では、発話文処理装置の一例として、従来例同様、異なる言語間の変換を行う通訳装置を用いて説明する。図1は本発明の請求項1から4に係る一実施の形態である通訳装置のブロック図である。

【0024】

実施の形態の通訳装置は、まず通訳する前に、形態素解析部2で対訳コーパス1内の原言語文の形態素解析を行うことで原言語文のみ品詞タグが付与された対訳コーパスを作成する。たとえば、図2の20の「部屋の予約をお願いしたいんですが」の発声文例では、21のような品詞タグが原言語文に与えられる。次に、品詞化部4で、上記コーパスの品詞タグ付き原言語文において、一部の単語名を品詞名に置きかえた品詞化対訳コーパスを作成する。この際に品詞名に変換される単語は以下の条件を満たすものとする。

- (1) 内容語テーブルに記載の品詞に対応する単語である。
- (2) 対訳単語辞書に登録されている単語で、対訳単語辞書の目的言語訳に相当する単語が、コーパス内の相当する目的言語対訳文に存在する。

【0025】

図2の内容語定義表3の例では、内容語テーブルに記載されている一般名詞、さ変名詞、動詞の中で、対訳単語辞書7に登録されている「部屋」と「予約」のみが品詞化され、22のようにこれらの単語を品詞名に置き換えたコーパスが作成される。さらに、相当する目的言語対訳文内の単語名も23のように日本語品詞名に置き換える。

【0026】

次に、上記の一部の内容語が品詞名に置き換えられたコーパスについて、フレーズ抽出部5は、原言語文、目的言語文別々に、各単語または品詞の2連鎖出現

頻度（以後 bi-gramと呼ぶ）を算出する。算出式を（数1）に示す。

【0027】

【数1】

単語（または品詞）i と単語（または品詞）j が隣接して出現した数
単語（または品詞）i の全出現数 + 単語（または品詞）j の全出現数

【0028】

コーパス内の全原言語文及び目的言語文を対象にbi-gramを算出した後、フレーズ抽出部5で、最も出現頻度の高かった2単語または品詞対を1つの単語とみなして連結し、再度bi-gramを算出する。これにより、たとえば頻度高く隣接する「お」「願い」、「願い」「し」、「し」「ます」などの単語対が連結され、「お願いします」というフレーズ候補が形成される。目的言語では「I'd」「like」、「like」「to」の単語対が連結される。全原言語文及び目的言語文別々に、以上の連結とbi-gram算出とを、bi-gramの値が全て一定閾値を超えなくなるまで繰り返す。そして、連結された単語も含めた個々の単語をフレーズ候補として抽出する。

【0029】

次にフレーズ決定部6で、原言語文と目的言語文対において、各フレーズが同時に出現している頻度を算出する。i番目の原言語フレーズをJ[i]、j番目の目的言語フレーズをE[j]とすると、フレーズJ[i]とE[j]との共起頻度K[i, j]は、算出式を（数2）にて算出される。

【0030】

【数2】

$$K[i, j] = \frac{\text{フレーズJ[i]とフレーズE[j]とが、対訳文対に共起する数}}{\text{フレーズJ[i]の出現数 + フレーズE[j]の出現数}}$$

【0031】

たとえば、図3の例では、原言語文の3つフレーズ30のうち、原言語フレーズの「お願いします」と目的言語フレーズの「I'd like to」との共起頻度は2/

(2+3)、「したいんですが」と目的言語フレーズの共起頻度は $1/(1+3)$ となる。この頻度が一定値以上のフレーズ対を対訳フレーズとして決定し、頻度と共にフレーズ番号を付けて対訳フレーズ辞書 9 に登録する。さらに、対訳フレーズとして決定されなかったフレーズ候補の中で、既に品詞化されている単語は、それ単独で対訳フレーズとして対訳フレーズ辞書 9 に登録する。それ以外の部分は、対訳対の中で各々の単語列どうしを一对としてフレーズ辞書に登録する。

【0032】

たとえば、図 3 の例では、31 のように対訳フレーズ辞書 9 に登録される。このようにして、フレーズ登録を行なった後、一文に共起するフレーズ番号を記録し、フレーズ番号対として対訳フレーズ間規則表 8 に登録する。図 3 の例では 32 となる。また、上記フレーズ番号対のフレーズ bi-gram を求め、これも対訳フレーズ間規則表 8 に記録する。例えば図 3 の 32 では、例えばフレーズ 3 とフレーズ 1 のフレーズ bi-gram を求める。またフレーズ 4、フレーズ 5、フレーズ 2 のフレーズ間規則に関してはフレーズ 4、フレーズ 5 及びフレーズ 5、フレーズ 2 の bi-gram をそれぞれ求め、対訳フレーズ間規則表 8 に記録する。

【0033】

通訳の際には、まず発声された原言語音声は音声認識部 10 に入力される。音声認識部では、たとえば、対訳フレーズ辞書 9 にフレーズとして記述されている単語のネットワークと対訳フレーズ間規則表 8 にて記述されているフレーズ bi-gram とにより、時系列に沿って順次認識単語候補が予測される。予め学習されている音響モデル 13 と入力音声との距離値をベースとした音響スコアとフレーズ bi-gram による言語スコアとの和を認識スコアとし、Nbest-search により認識候補である連続単語列が決定される。

【0034】

認識された連続単語列は、言語変換部 11 に入力される。言語変換部 11 では、入力された連続単語列を対訳フレーズ辞書 9 内のフレーズ列に変換し、各フレーズ列に相当するフレーズ間規則を探索する。そして、各フレーズの対訳である目的言語フレーズと目的言語のフレーズ間規則とから、入力原言語認識結果文を

目的言語文に変換する。

【0035】

変換された目的言語文は出力文生成部 12 に入力され、統語的な不自然さを修正する。たとえば、定冠詞や不定冠詞の付与、代名詞、動詞における 3 人称化や複数化や過去形化などの最適化などが行われる。修正後の目的言語翻訳結果文はたとえばテキストとして出力される。この動作は従来例と同様である。

【0036】

以上の実施の形態では、原言語文及び目的言語文各々における単語または品詞の隣接頻度と、対訳における頻度の高い単語列または品詞列の共起関係を用いて自動的に対訳フレーズとフレーズ間規則を決定し、この対訳フレーズ規則を用いて言語または文体変換とを行うことにより、効率よく品質の高い対訳フレーズ辞書を生成できる言語変換装置を提供することが出来る。

【0037】

(実施の形態 2)

次に請求項 5 における実施の形態を説明する。本実施の形態も、言語変換装置の一例として、第 1 の実施の形態同様、異なる言語間の変換を行う通訳装置を用いて説明する。図 4 は本発明の請求項 5 に係る一実施の形態である通訳装置のブロック図である。

【0038】

実施の形態の通訳装置は、まず通訳する前に、第 1 の実施の形態同様、形態素解析後、品詞タグが付与された対訳コーパスを作成する。

【0039】

次に、フレーズ抽出部 42 で、予めフレーズとして抽出したい単語または品詞列を規則化して記述してあるフレーズ定義表 41 に従い、規則に相当する単語または品詞を連結する。たとえば図 5 の 41 の例では、「動詞+助動詞」や「格助詞+動詞」などの規則により、「を+(動詞)+たい」が単語として連結される。このように、上記の一部の内容語が品詞名に置き換えられ、さらに上記のような単語または品詞列が連結され一単語とみなされたコーパスについて、原言語文、目的言語文別々に、各単語または品詞の 2 連鎖出現頻度 (以後 bi-gram と呼ぶ)

を算出する。算出式は（数 1）と同様である。

【0040】

さらに、bi-gramの値が全て一定閾値を超えなくなるまで、第 1 の実施の形態と同等に、処理を繰り返す。そして、連結された単語も含めた個々の単語をフレーズ候補として抽出し、フレーズ決定部で、第 1 の実施の形態と同様に対訳フレーズ辞書 44 と対訳フレーズ間規則表 45 を作成する。

【0041】

通訳の際の動作も第 1 の実施の形態と同様である。

【0042】

以上の実施の形態では、予め定義されているフレーズとみなしたい単語または品詞列の規則に従って単語または品詞を連結した他後、原言語文及び目的言語文各々における単語または品詞の隣接頻度と、対訳における頻度の高い単語列または品詞列の共起関係を用いて自動的に対訳フレーズとフレーズ間規則を決定し、この対訳フレーズ規則を用いて言語または文体変換とを行うことにより、人手を最小限度に押さえた範囲で、さらに効率よく品質の高い対訳フレーズ辞書を生成できる言語変換装置を提供することが出来る。

【0043】

なお、本実施の形態の対訳フレーズは、本発明の対応するフレーズの例である。

【0044】

さらに、本発明の言語変換装置の各構成要素の全部または一部の機能を専用のハードウェアを用いて実現しても構わないし、またコンピュータのプログラムによってソフトウェア的に実現しても構わない。

【0045】

さらに、本発明の言語変換装置の各構成要素の全部または一部の機能をコンピュータに実行させるためのプログラムを格納していることを特徴とするプログラム記録媒体も本発明に属する。

【0046】

【発明の効果】

以上説明したところから明らかなように、本発明に係る請求項 1 記載の言語変換装置は、隣接する頻度の高い単語または品詞列を連結してフレーズを生成し、原言語文及び目的言語文を対応させながら対訳フレーズを決定することにより、コーパスから対訳フレーズを自動的に抽出することを可能とし、人手をなるべく用いずに、効率よく品質の高い対訳フレーズ辞書を生成できる。

【0047】

また、請求項 2 に記載の言語変換装置は、対訳フレーズを決定する際に、原言語及び目的言語のフレーズの共起頻度を調べることで、コーパスから対訳フレーズを自動的に抽出することを可能とし、人手をなるべく用いずに、効率よく品質の高い対訳フレーズ辞書を生成できる。

【0048】

また、請求項 3 及び 4 に記載の言語変換装置においては、予め対訳コーパスの一部の単語を品詞化しておき、品詞化された対訳コーパスから対訳フレーズを抽出することで、自動的にフレーズを抽出する際に、コーパスに含まれる単語以外の単語が含まれているフレーズにも対応できるため、より効率よく品質の高い対訳フレーズ辞書を生成できる。

【0049】

また、請求項 5 に記載の言語変換装置は、請求項 1 の動作で自動抽出される対訳フレーズと、予め人為的に定義された対訳フレーズとの両方を対訳フレーズ辞書に登録することにより、人手を最小限度に押さえた範囲で、より品質の高い対訳フレーズ辞書を生成できる。

【0050】

【図面の簡単な説明】

【図 1】

本発明の第 1 の実施の形態における言語変換装置の構成を示すブロック図

【図 2】

本発明の第 1 の実施の形態における発声文例から品詞化対訳コーパスを作成する例を示す図

【図 3】

本発明の第 1 の実施の形態における対訳フレーズ辞書と対訳フレーズ間規則表の例を示す図

【図 4】

本発明の第 2 の実施の形態における言語変換装置の構成を示すブロック図

【図 5】

本発明の第 2 の実施の形態におけるフレーズ定義表に従い規則に相当する単語または品詞を連結した例を示す図

【図 6】

従来 of 言語変換装置 of ブロック図

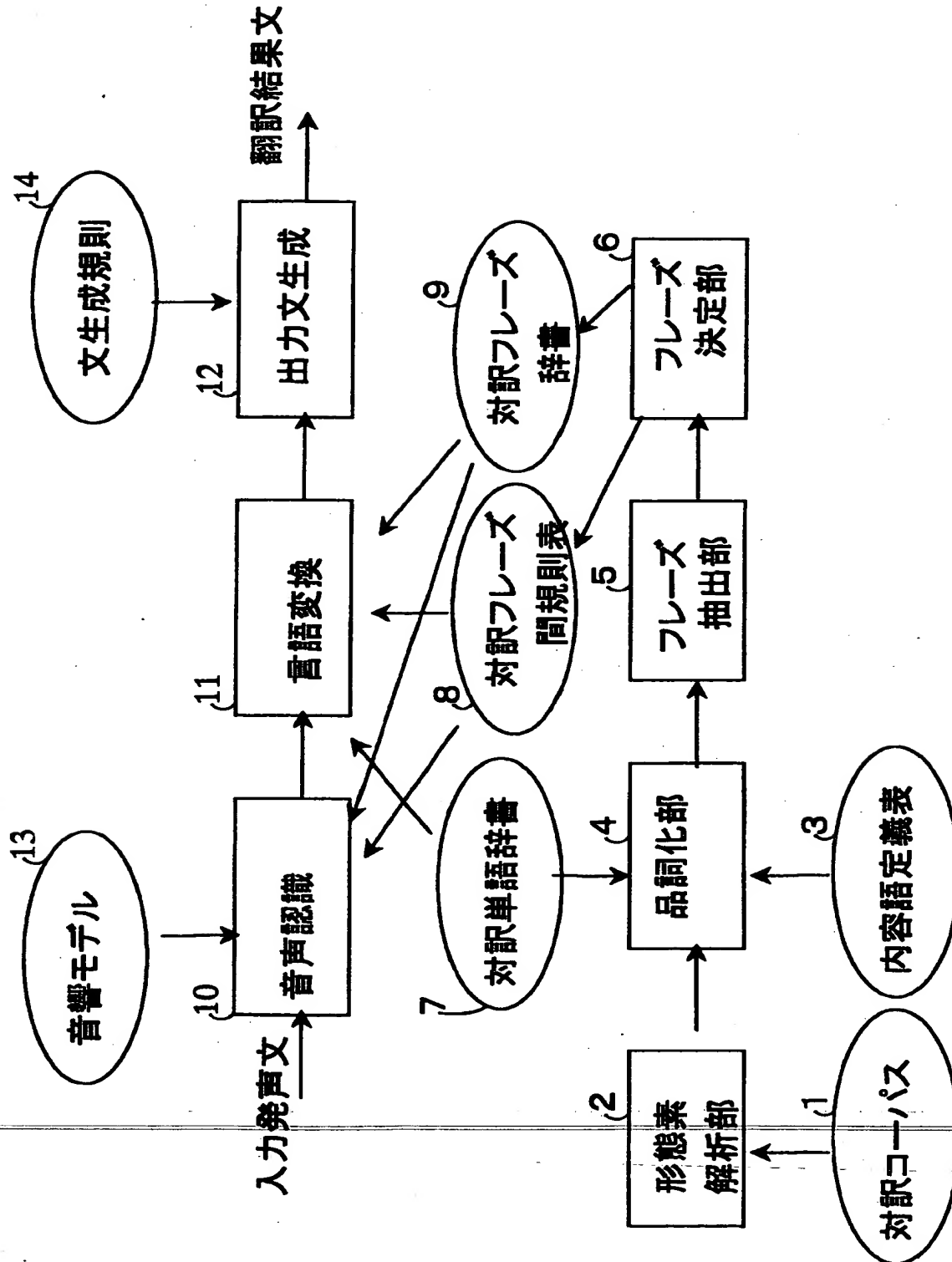
【符号 of 説明】

- 1 : 対訳コーパス
- 2 : 形態素解析部
- 3 : 内容語定義表
- 4 : 品詞化部
- 5、4 2 : フレーズ抽出部
- 6、4 3 : フレーズ決定部
- 7 : 対訳単語辞書
- 8、4 5 : フレーズ間規則表
- 9、4 6 : 対訳フレーズ辞書
- 1 0 : 音声認識部
- 1 1 : 言語変換部
- 1 2 : 出力文生成部
- 1 3 : 音響モデル
- 1 4 : 文生成規則

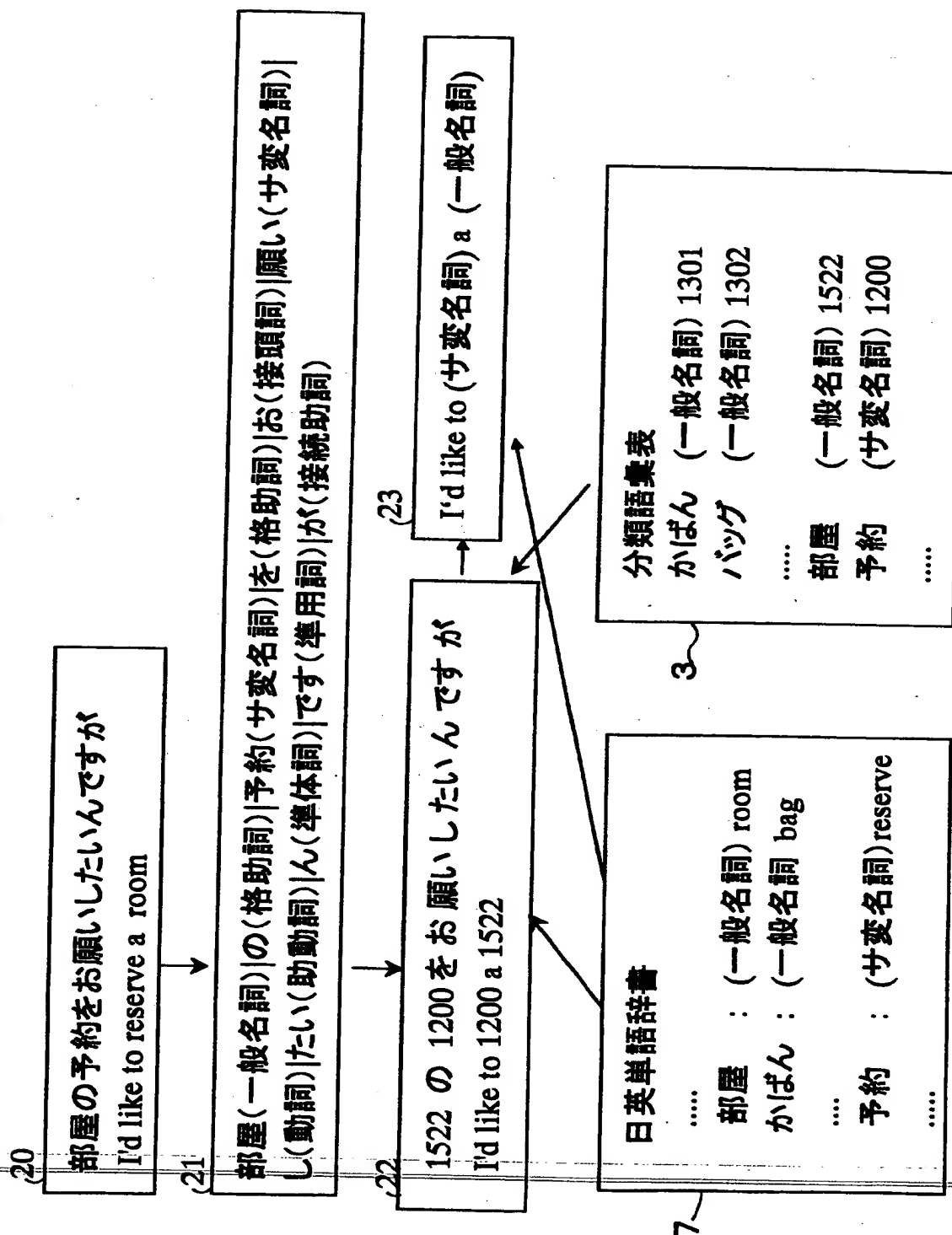
【書類名】

図面

【図 1】



【図 2】



【図 3】

30

[対訳文1]: (一般名詞) の (サ変名詞) | お願いします

I'd like to | (一般名詞) a (一般名詞)

[対訳文2]: (一般名詞) | お願いします

I'd like to | (一般名詞)

[対訳文3]: (一般名詞) | を | (動詞) | たいんですが

I'd like to | (動詞) | (一般名詞)

31

[フレーズ1]: お願いします : I'd like to | 2 / (3+2)

[フレーズ2]: たいんですが : I'd like to | 1 / (3+1)

[フレーズ3]: (一般名詞) の (サ変名詞) | (一般名詞) a (一般名詞) |

[フレーズ4]: (一般名詞) : (一般名詞) |

[フレーズ5]: (動詞) : (動詞) |

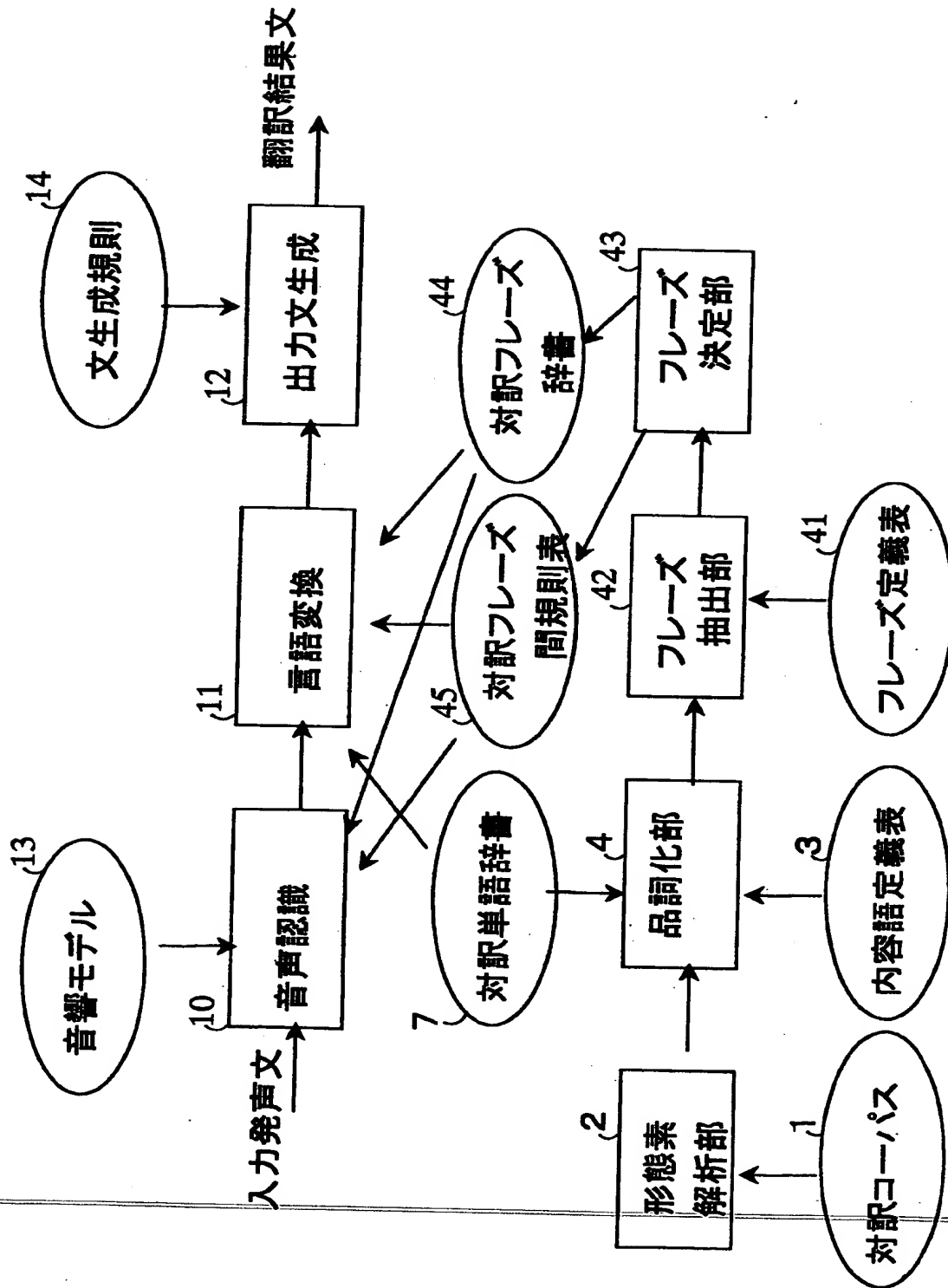
32

[フレーズ3] [フレーズ1] : [フレーズ1] [フレーズ3]

[フレーズ4] [フレーズ1] : [フレーズ4] [フレーズ1]

[フレーズ4] [フレーズ5] [フレーズ2] : [フレーズ2] [フレーズ4]

【図4】



【図 5】

41

格助詞＋動詞または動詞を含むフレーズ
動詞＋助動詞または助動詞を含むフレーズ
形容詞＋名詞
.....

51

[対訳文1]: (一般名詞) の (サ変名詞) | お願いします

I'd like to | (サ変名詞) a (一般名詞)

[対訳文2]: (一般名詞) | お願いします

I'd like to | (一般名詞)

[対訳文3]: (一般名詞) | を (動詞) たいんですが

I'd like to (動詞) | (一般名詞)

52

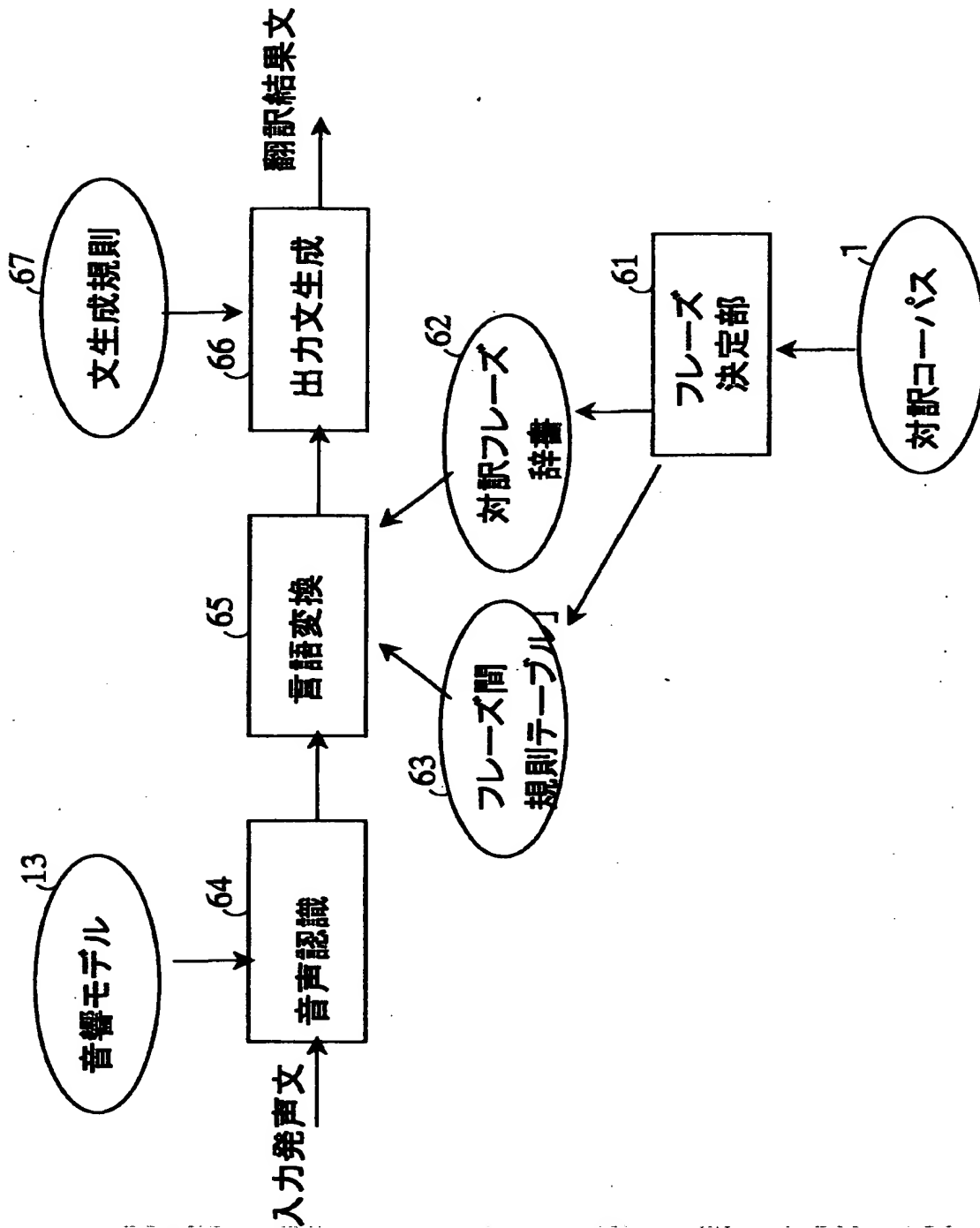
[フレーズ1]: お願いします : I'd like to | 2 / (3+2)

[フレーズ2]: を (動詞) たいんですが : I'd like to (動詞) | 1 / (3+1)

[フレーズ3]: (一般名詞) の (サ変名詞) | (サ変名詞) a (一般名詞) |

[フレーズ4]: (一般名詞) : (一般名詞) |

【図6】



【書類名】 要約書

【要約】

【課題】 対訳フレーズ辞書やフレーズ間規則は、発声文の意味的情報や文法的情報を自動的に解析できる手段がないために、人手で作成しなければならない。そのため、開発に時間がかかり、人手の作成基準の揺れが規則性能を歪ませる。また、通訳装置の目標となるタスクを変更したり、原言語及び目的言語の言語種が変更になった場合は、一度構築した規則を適応できずにはじめから規則を作成しなければならず、開発効率が悪く手間がかかる。

【解決手段】 隣接する頻度の高い単語または品詞列を連結してフレーズを生成するフレーズ抽出部 5 と、原言語文及び目的言語文を対応させながら対訳フレーズを決定するフレーズ決定部 6 とを備え、コーパスから対訳フレーズを自動的に抽出することを可能とし、人手をなるべく用いずに、効率よく品質の高い対訳フレーズ辞書 9 を生成できる。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000005821]

1. 変更年月日	1990年 8月28日
[変更理由]	新規登録
住 所	大阪府門真市大字門真1006番地
氏 名	松下電器産業株式会社

This Page Blank (uspto)